AMENDMENTS TO THE SPECIFICATION

Please replace originally filed paragraphs [0001], [0017], [0020], [0054], [0057], [0058], [0063], [0074], [0097], [0150], [0155], [0160], [0162], [0168], [0184], [0187] and [0193] with the following amended replacement paragraph(s). Deletions of original text are indicated with ~~strikethrough~~ or double bracket [[xx]]. Added text is shown by underline.

[0017] In one embodiment, an *E. coli* cell includes an orthogonal leucyl-tRNA (leucyl-O-tRNA), where the leucyl-O-tRNA comprises at least about a 25% suppression activity in presence of a cognate synthetase in response to a selector codon as compared to a control lacking the selector codon; and an orthogonal leucyl aminoacyl-tRNA synthetase (leucyl-O-RS), where the O-RS preferentially aminoacylates the O-tRNA with a selected amino acid. The E. coli cell also includes the selected amino acid, and, a nucleic acid that comprises a polynucleotide that encodes a polypeptide of interest, where the polynucleotide comprises a selector codon that is recognized by the leucyl O-tRNA. In one example, the leucyl O-tRNA is derived from *Halobacterium sp NRC-1* and the leucyl O-RS is derived from *Methanobacterium thermoautotrophicum* ~~*thermoaautotropicum*~~.

[0019] The leucyl O-tRNA and/or the leucyl O-RS of the invention can be derived from any of a variety of organisms (e.g., both eukaryotic and non-eukaryotic organisms). For example, the leucyl O-tRNA is derived from an archael tRNA (e.g., from *Halobacterium sp NRC-1*) and/or the leucyl O-RS is derived from a non-eukaryotic organism (e.g., *Methanobacterium thermoautotrophicum* ~~*thermoaautotropicum*~~).

[0020] Polynucleotides are also a feature of the invention. A polynucleotide of the invention includes a polynucleotide comprising a nucleotide sequence as set forth in any one of SEQ ID NO.: 1-2, 4-7, 12, and/or is complementary to or that encodes a polynucleotide sequence of the above. A polynucleotide of the invention also includes a nucleic acid that hybridizes to a polynucleotide described above, under highly stringent conditions over substantially the entire length of the nucleic acid. A polynucleotide of the invention also includes a polynucleotide that is, e.g., at least 75%, at least 80%, at least 90%, at least 95%, at least 98% or more identical to that of a naturally occurring leucyl tRNA or a consensus sequence of multiple naturally

occurring leucyl tRNAs, e.g., the leucyl ~~laucyl~~ tRNA of SEQ ID NO: 12, and comprises an anticodon loop comprising a $CU(X)_n$ XXXAA sequence, an stem region lacking noncanonical base pairs and a conserved discriminator base at position 73. A polynucleotide of the invention also includes a polynucleotide that is, e.g., at least 75%, at least 80%, at least 90%, at least 95%, at least 98% or more identical to that of a naturally occurring leucyl tRNA and comprises an anticodon loop comprising a CUUCCUAA sequence, a first pair selected from T28:A42, G28:C42 and/or C28:G42, and a second pair selected from G:49:C65 or C49:G65, where the numbering corresponds to that indicated in Figure 4, Panel A. Polynucleotides that are, e.g., at least 80%, at least 90%, at least 95%, at least 98% or more identical to any of the above and/or a polynucleotide comprising a conservative variation of any the above or in Table 3 are also polynucleotides of the invention.

[0054] **Figure 1, Panels A, B** and **C** schematically illustrate leucyl tRNAs and synthetases, and their phylogenetic relationships. **Panel A** provides a ClustalW analysis of aminoacyl-tRNA synthetases, where Archaeal tRNA synthetases are labeled using a dashed line, prokaryotic using a solid line, and eukaryotic sequences using a dotted line. This analysis reveals the halobacterial synthetase to be unusual in its homology to prokaryotic rather than archaeal and eukaryotic synthetases. **Panel B** provides a ClustalW analysis of Halobacterial tRNAs which all share high homology to other archaeal tRNAs. Dendrograms were generated using the program PhyloDraw. **Panel C** provides a sequence alignment of multiple sequences of the family of archaeal leucyl tRNAs examined as potential orthogonal suppressors (SEQ ID NOs: 37 (top of figure) through 72 (bottom of figure)). Sequences examined as potential amber suppressors by changing the anticodon (boxed) to CUA are shown in bold as is the consensus sequence. The highly conserved positions G37 and A73 are indicated with underlining.

[0057] **Figure 4, Panels A** and **B** illustrates the optimization of suppressor tRNAs. **Panel A** illustrates regions (shown in boxes) of the halobacterial orthogonal tRNA (SEQ ID NO: 11) subjected to mutagenesis in an effort to improve the efficiency or selectivity of TAG and AGGA suppressor tRNAs. **Panel B** illustrates that active mutant TAG suppressors identified by positive selection conserve A73. Less cross-reactive mutants identified by a double-sieve selection strategy conserve a C3:G70 base pair. The most active and selective suppressor tRNA is shown with double boxes.

[0058] **Figure 5** illustrates a consensus-derived frameshift suppressor (SEQ ID NO: 12). A consensus sequence was obtained by multiple sequence alignment of all known archaeal leucyl tRNAs, and the anticodon loop is changed to UCUCCUAA. The variations observed for tRNAs identified by selection are shown in boxes. The most active mutations are shown with double boxes.

[0063] Translation systems that are suitable for making proteins that include one or more selected amino acids, e.g., an unnatural amino acid, are described in International patent applications WO 2002/086075, entitled "METHODS AND COMPOSITION FOR THE PRODUCTION OF ORTHOGANOL tRNA-AMINOACYLtRNA SYNTHETASE PAIRS" and WO 2002/085923, entitled "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS." In addition, see WO 2004/094593 entitled "EXPANDING THE EUKARYOTIC GENETIC CODE," ~~International Application Number PCT/US2004/011786,~~ filed April 16, 2004. Each of these applications is incorporated herein by reference in its entirety. These translation systems can be adapted to the present invention by substituting the leucyl-O-RS and leucyl-O-tRNA provided herein.

[0074] A leucyl O-tRNA of the invention comprises at least about a 25% suppression activity in presence of a cognate synthetase in response to a selector codon, as compared to a control lacking the cognate synthetase. Suppression activity can be determined by any of a number of assays known in the art. For example, a β-galactosidase reporter assay can be used. A derivative of a plasmid that expresses *lacZ* gene under the control of promoter is used, e.g., where the Leu-25 of the peptide VVLQRRDWEN (SEQ ID NO: 22) of *lacZ* is replaced by a selector codon, e.g., TAG, TGA, AGGA, etc. codons, or sense codons (as a control) for tyrosine, serine, leucine, etc. The derivatived lacZ plasmid is introduced into cells from an appropriate organism (e.g., an organism where the orthogonal components can be used) along with plasmid comprising a O-tRNA of the invention. A cognate synthetase can also be introduced (either as a polypeptide or a polynucleotide that encodes the cognate synthetase when expressed). The cells are grown in media to a desired density, e.g., to an OD_{600} of about 0.5., and β-galactosidase assays are performed, e.g., using the BetaFluor™ β-Galactosidase Assay Kit (Novagen). Percent suppression is calculated as the percentage of activity for a sample relative to a

comparable control, e.g., the value observed from the derivatived lacZ construct, where the construct has a corresponding sense codon at desired position rather than a selector codon.

[0097] Additional general details for producing O-RS, and altering the substrate specificity of the synthetase can be found in WO 2002/086075 entitled "Methods and compositions for the production of orthogonal tRNA-aminoacyltRNA synthetase pairs; " and WO 2004/094593 entitled "EXPANDING THE EUKARYOTIC GENETIC CODE," ~~International Application Number PCT/US2004/011786,~~ filed April 16, 2004.

[0150] One example of an algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul *et al.*, J. Mol. Biol. 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information ~~(www.ncbi.nlm.nih.gov/)~~. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915).

[0155] A catalogue of Bacteria and Bacteriophages useful for cloning is provided, e.g., by the ATCC, e.g., *The ATCC Catalogue of Bacteria and Bacteriophage* (1992) Gherna *et al.* (eds) published by the ATCC. Additional basic procedures for sequencing, cloning and other aspects of molecular biology and underlying theoretical considerations are also found in Watson *et al.* (1992) *Recombinant DNA Second Edition Scientific American Books, NY.* In addition, essentially any nucleic acid (and virtually any labeled nucleic acid, whether standard or non-standard) can be custom or standard ordered from any of a variety of commercial sources, such as the Midland Certified Reagent Company (Midland, TX ~~mcre.com~~), The Great American Gene Company (Ramona, CA ~~available on the World Wide Web at genco.com~~), ExpressGen Inc. (Chicago, IL ~~available on the World Wide Web at expressgen.com~~), Operon Technologies Inc. (Alameda, CA) and many others.

[0160] The compositions of the invention and compositions made by the methods of the invention optionally are in a cell. The leucyl O-tRNA/O-RS pairs or individual components of the invention can then be used in a host system's translation machinery, which results in a selected amino acid, e.g., unnatural amino acid, being incorporated into a protein. WO 2004/094593 entitled "EXPANDING THE EUKARYOTIC GENETIC CODE," ~~The International Application Number PCT/US2004/011786,~~ filed April 16, 2004, and WO 2002/085923, entitled "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS" describe this process, and is incorporated herein by reference. For example, when an leucyl O-tRNA/O-RS pair is introduced into a host, e.g., *Escherichia coli*, the pair leads to the in vivo incorporation of selected amino acid, such as an unnatural amino acid, e.g., a synthetic amino acid, such as derivative of a leucine amino acid, which can be exogenously added to the growth medium, into a protein, in response to a selector codon. Optionally, the compositions of the present invention can be in an in vitro translation system, or in an in vivo system(s).

[0162] Typically, the proteins are, e.g., at least 60%, at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, or at least 99% or more identical to any available protein (e.g., a therapeutic protein, a diagnostic protein, an industrial enzyme, or portion thereof, and the like), and they comprise one or more selected amino acid. Examples of therapeutic, diagnostic, and other proteins that can be modified to comprise one or more selected amino acid, e.g., an unnatural amino acid, can be found, but not limited to, those in WO 2004/094593 entitled

"EXPANDING THE EUKARYOTIC GENETIC CODE," ~~International Application Number PCT/US2004/011786,~~ filed April 16, 2004, and WO 2002/085923, entitled "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS."

[0168] Additional details on proteins, antibodies, antisera, etc. can be found in, e.g., WO 2004/094593 entitled "EXPANDING THE EUKARYOTIC GENETIC CODE," ~~International Application Number PCT/US2004/011786~~, filed April 16, 2004, WO 2002/085923, entitled "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS;" WO 2004/035605 ~~International Application Numbers PCT/US2003/32870~~, filed October 15, 2003; and WO 2004/058946 ~~PCT/US2003/41346~~, filed December 22, 2003.

[0184] *Beta-galactosidase reporter assays.* The full-length *lacZ* gene of plasmid pBAD-Myc/His/lacZ (Invitrogen) was amplified by PCR and subcloned into plasmid pLASC to obtain plasmid pLASC-lacZ. This pSC101-derived plasmid expresses *lacZ* gene under the control of an *lpp* promoter and has an ampicillin resistance gene for plasmid maintenance. Derivatives of this plasmid were constructed wherein Leu-25 of the peptide VVLQRRDWEN (SEQ ID NO: 22) of *lacZ* was replaced by TAG, TGA, or AGGA codons, or sense codons for tyrosine, serine, or leucine. The appropriate pLASC-lacZ-, pACKO-Bla-, and pKQ-derived plasmids were cotransformed and grown to an OD_{600} of 0.5. Beta-galactosidase assays were performed in quadruplicate using the BetaFluor™ β-Galactosidase Assay Kit (Novagen). Percent suppression was calculated as the percentage of activity for a sample relative to the value observed from the pLASC-lacZ construct with the corresponding sense codon at position 25. Cells containing pLASC-lacZ plasmids with sense codons at position 25 were also assayed by 2-nitrophenyl-β-D-galactopyranoside assays (Miller (1972) Experiments in molecular genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), and activity was calculated in Miller units.

[0187] Detailed information for the cloning of archaeal leucyl-tRNA synthetases can be found below and is also available on the American Chemical Society publications website ~~Internet at http://pubs.acs.org~~.

TABLE 2: CLONING OF ARCHAEAL LEUCYL-tRNA SYNTHETASES INTO PLASMID pKQ

Organism	Accession number	ATCC Number	Forward Oligo	Reverse Oligo
Halobacterium sp. NRC-1	NP_280869.1	700922	ca214	ca215

Escherichia coli (strain HB101)	P07813	N/A	ca244	ca215
Methanococcus jannaschii	Q58050	43067D	ca246	ca247
Archaeoglobus fulgidus	O30250	49558D	ca261	ca247
Aeropyrum pernix K1	Q9YD97	700793D	ca263	ca264
Pyrococcus horikoshii	O58698	700860	ca265	ca266
Methanobacterium thermoautotrophicum	O27552	700791	ca274	ca275

Organism	Restriction enzymes
Halobacterium sp. NRC-1	NcoI/EcoRI
Escherichia coli (strain HB101)	NcoI/EcoRI
Methanococcus jannaschii	NcoI/KpnI
Archaeoglobus fulgidus	NcoI/KpnI
Aeropyrum pernix K1	BspHI/EcoRI (subcloned into NcoI/EcoRI sites)
Pyrococcus horikoshii	NcoI/EcoRV (subcloned into NcoI/PvuII sites)
Methanobacterium thermoautotrophicum	BsmBI (subcloned into NcoI/EcoRI sites)

Oligo Sequences

ca214	GGTTTCCATGGGAGAGCAAGCCACCTAC	SEQ ID NO: 23
ca215	GGTTTGGAATTCAGTCGTCGGCTTCGTCG	SEQ ID NO: 24
ca244	CGAAACCATGGAAGAGCAATACCGCCCGGAAG	SEQ ID NO: 25
ca245	CCAAAGAATTCCCGCCAACGACCAGATTGAGGAG	SEQ ID NO: 26
ca246	CGAAACCATGGTTATGATTGACTTTAAAG	SEQ ID NO: 27
ca247	CGAAAGGTACCTTGTATTCAAGATAAATAGCTGG	SEQ ID NO: 28
ca261	GCGAACCATGGGCGATTTCAGGATAATTGAG	SEQ ID NO: 29
ca262	CAATTGGTACCTTAAGCAACATAAATCGCG	SEQ ID NO: 30
ca263	GGATTATCATGAAGCGACTAAAGGCCGTGGAGGAG	SEQ ID NO: 31
ca264	CACTTGAATTCTTAGCCTCCTCTCTTCTCCGC	SEQ ID NO: 32
ca265	CGAATCCATGGCTGAGCTTAACTTCAAGG	SEQ ID NO: 33
ca266	GGATGGATATCACTCGATGAAGATGGCAG	SEQ ID NO: 34
ca274	GGAGACGTCTCTCATGGATATTGAAAGAAAATGGCG	SEQ ID NO: 35
ca275	CGTTACGTCTCGAATTGGAAAAGAGCTGTCTGAGG	SEQ ID NO: 36

[0193] *Optimization of the tRNA anticodon loop.* The robust endogenous amber suppressor *supD* confers survival to 1000 μg/mL ampicillin when expressed from pACKO-A184TAG. In contrast, cells expressing the MtLRS/HL(TAG)1 pair survive to only 35 μg/mL ampicillin, which corresponds to a 2.9% suppression efficiency as determined from β-galactosidase assays (Table 1). We therefore sought to improve the activity of the system. Previous experiments on frameshift, missense, and nonsense suppression revealed that A37 was a highly conserved feature in robust suppressor tRNAs (Magliery et al., (2001) J. Mol. Biol. 307, 755-769). HhL4 has a G at position 37, therefore substitution of G37 to A might be expected to

improve suppression efficiency. To examine this and other possible anticodon loop mutants, a library was constructed in which the 7 positions of the anticodon loop (positions 32-38, see **Figure 4, Panel A**) in HhL4 were replaced with degenerate bases and subcloned into pACKO-A184TAG. The library of tRNAs was cotransformed with pKQ-MtLRS and subjected to ampicillin selection initially at 35 μg/mL ampicillin for two rounds of selection, then plated on a series of plates with increasing ampicillin concentration in the third round of selection. At the highest concentration of ampicillin for which growth was observed (500 μg/mL), the only clone found had an anticodon loop with the sequence CUCUAAA, corresponding to a simple G37A mutation (Table 1). When cotransformed with pKQ-MtLRS, this clone could survive to 500 μg/mL ampicillin. In the absence of the synthetase it survived to only 25 μg/mL ampicillin. Under similar conditions, cells containing the wild-type *M. jannaschii* tyrosyl orthogonal amber suppressor tRNA survive to 350 μg/mL ampicillin in the presence of the cognate synthetase and to 60 ug/mL ampicillin without the synthetase.

Table 1. Suppression efficiency of mutant orthogonal tRNAs (SEQ ID NOs: 1-10)

Reporter Plasmid	Miller Units
pLASC-lacZ(Leu)	210 ± 2
pLASC-lacZ(Ser)	200 ± 5
pLASC-lacZ(Tyr)	192 ± 7
pLASC-lacZ(TAG)	1 ± 1
pLASC-lacZ(AGGA)	2 ± 1
pLASC-lacZ(TGA)	1 ± 1

	Percent Suppression[a]			
Suppressor tRNA	with pKQ	with synthetase	Sequence	SEQ ID NO
HL(TAG)1	0.4 ± 0.1 %	2.9 ± 0.8 %	GCGAGGGTAGCCAAGCTCGGCCAACGGCGACGGACT**CTAG**ATCCGTT CTCGTAGGAGTTCGAGGGTTCGAATCCCTTCCCTCGCACCA	1
HL(TAG)2	0.3 ± 0.1 %	9.6 ± 0.4 %	GCGAGGGTAGCCAAGCTCGGCCAACGGCGACGGACTCTA**A**ATCCGTT CTCGTAGGAGTTCGAGGGTTCGAATCCCTTCCCTCGCACCA	2
HL(TAG)3	1.5 ± 1.2 %	33.2 ± 4.4 %	CCCAGGGTAGCCAAGCTCGGCCAACGGCGACGGACTCTAAATCCGTT CTCGTAGGAGTTCGAGGGTTCGAATCCCTTCCCT**G**GGACCA	3
HL(AGGA)1	0.4 ± 0.1 %	4.6 ± 2.1 %	GCGAGGGTAGCCAAGCTCGGCCAACGGCGACGGACT**TCCTA**ATCCGT TCTCGTAGGAGTTCGAGGGTTCGAATCCCTTCCCTCGCACCA	4
HL(AGGA)2	0.7 ± 0.3 %	14.9 ± 6.1 %	GCGAGGGTAGCCAAGCTCGGCCAACGGCGACGGACTTCCTAATCCGT TCTCGTAGGAGTTCGAGGGTTCGAATCCCT**C**CCCTCGCACCA	5
HL(AGGA)3	7.4 ± 0.4 %	35.5 ± 1.4 %	GCGGGGGTTGCCGAGCCTGGCCAAAGGCGCCGGACTTCCTAATCC**G**G TCCCGTAGGGGTTCCGGGGTTCAAATCCCCGCCCCCGCACCA	6
HL(TGA)1	4.7 ± 1.5 %	60.8 ± 7.0 %	GCGGGGGTTGCCGAGCCTGGCCAAAGGCGCCGGACT**TCAA**ATCCGGT CCCGTAGGGGTTCCGGGGTTCAAATCCCCGCCCCCGCACCA	7
J17[b]	0.2 ± 0.1 %	18.5 ± 4.8 %	CCGGCGGTAGTTCAGCAGGGCAGAACGGCGGACTCTAAATCCGCATG GCGCTGGTTCAAATCCGGCCCGCCGGACCA	8

SupD	42.8 ± 7.1 %	ND	GGAGAGATGCCGGAGCGGCTGAACGGACCGGTCTCTAAAACCGGAGT AGGGGCAACTCTACCGGGGGTTCAAATCCCCCTCTCTCCGCCA	9
Ser2AGGA	25.2 ± 0.1 %	ND	GGAGAGATGCCGGAGCGGCTGAACGGACCGGTCTTCCTAAACCGGAG TAGGGGCAACTCTACCGGGGGTTCAAATCCCCCTCTCTCCGCCA	10

[a]β-Galactosidase activity was determined for tRNA reporter plasmids derived from pACKO-Bla cotransformed with the appropriate pLASC-lacZ mutant and either a synthetase-expressing plasmid or a plasmid with no synthetase. Activity is reported as the percentage of activity observed relative to the value observed from the pLASC-lacZ construct with a leucyl (wild-type), seryl, or tyrosyl sense codon at position 25. In each case, the codon at position 25 of *lacZ* is designated in parentheses.

[b]J17, the *M. jannaschii* tyrosyl amber suppressor tRNA with improved orthogonality (Wang and Schultz (2001) Chem. Biol. 8, 883-890) was expressed in plasmid pACKO-A184TAG in the presence of pLASC-lacZ(TAG) and either pKQ or pBK-JYRS.